Osteoarthritis and Cartilage 20 (2012) 1470–1476

Osteoarthritis and Cartilage

OARSI OSTEOARTHRITIS RESEARCH SOCIETY INTERNATIONAL

Review

# Osteoarthritis year 2012 in review: genetics and genomics

J.B.J. van Meurs †‡*, A.G. Uitterlinden †‡§

† Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands
‡ The Netherlands Genomics Initiative-sponsored Netherlands Consortium for Healthy Aging (NGI-NCHA), Leiden/Rotterdam, The Netherlands
§ Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands

ARTICLE INFO

*Article history:*
Received 7 June 2012
Accepted 8 August 2012

*Keywords:*
Genetics
Genomics
DNA-variants
Osteoarthritis

SUMMARY

The field of genetics and genomics is a highly technological driven field that is advancing fast. The purpose of this year in review of genetics and genomics was to highlight the publications that apply these new technologies tools to improve understanding of the pathophysiology of osteoarthritis (OA). In addition, most recent developments in genetics and genomics research and their relevance to OA are discussed in this review.



## Introduction

Failure to understand the pathophysiology of osteoarthritis (OA) frustrates efforts to improve therapeutic strategies. Genome-wide genetic and genomic approaches have the potential to find novel biological pathways involved in OA, since these methods are hypothesis-free and do not suffer from the bias of previous knowledge. The discovery of novel genes might help to understand the development of OA and has the potential to identify novel treatment options. In addition, more personalized medicine options for OA can be explored through prediction of risk for disease as well as classification in disease subtypes.

Over the last year, substantial advances have been made in genetics and genomics of OA as evidenced by results from the TREAT-OA and arcOGEN consortia. Powerful genome-wide association studies (GWAS) have been performed and new genomics technology have been applied to the field over the last year, which are now beginning to bear fruit. This review is a summary of studies, selected by the author, related to genetics and genomics published between May 2011 and May 2012 and on data presented during the OA Research Society International (OARSI) 2012, which was held from April 26–29 2012 in Barcelona, Spain. In addition, most recent developments in genetics and genomics research and their relevance to OA will be discussed.

Genetic research is done (1) to identify underlying causative genes and pathways, and thereby understand more about the biology of the disease with potential implications for development of novel treatments, and (2) to be able to predict risk of disease by genotyping the identified risk alleles. So far, most research is focused on identifying novel OA genes to understand more of the disease biology, resulting in the identification of common genetic variants with modest effect size.

## Genetic architecture of OA

Through twin studies it has been well-established that OA and its endophenotypes, are to a large extent genetically determined, but the underlying genetic variants are mostly unknown. In terms of how many, what type, and which genetic variants explain the genetic variance of OA, i.e., the genetic architecture, we can distinguish several possibilities. The genetic architecture of OA is similar to other complex diseases with contributions of several and perhaps hundreds of genes, with most having small effects and a few having large effects (Fig. 1). Among these we can distinguish early onset OA, which usually represents a monogenic Mendelian disease type that can be mapped by linkage analysis in families, or nowadays by exome sequencing of affected subjects. On the other hand we can recognize late onset OA, which represents the common form of OA with a usual age at onset of >60 years, and for which genetic association approaches have been shown to be fruitful in identifying underlying genetic factors.

### *Early onset OA*

The early onset syndromic OA has an obvious genetic defect and high penetrance in a low number of families. These syndromes

* Address correspondence and reprint requests to: J.B.J. van Meurs, Genetic Laboratory, Department of Internal Medicine, Room Ee579b, Erasmus MC, PO Box 2040, 3000 CA, Rotterdam, The Netherlands. Tel: 31-107038425.
*E-mail address:* j.vanmeurs@erasmusmc.nl (J.B.J. van Meurs).

1063-4584/$ – see front matter © 2012 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.joca.2012.08.007

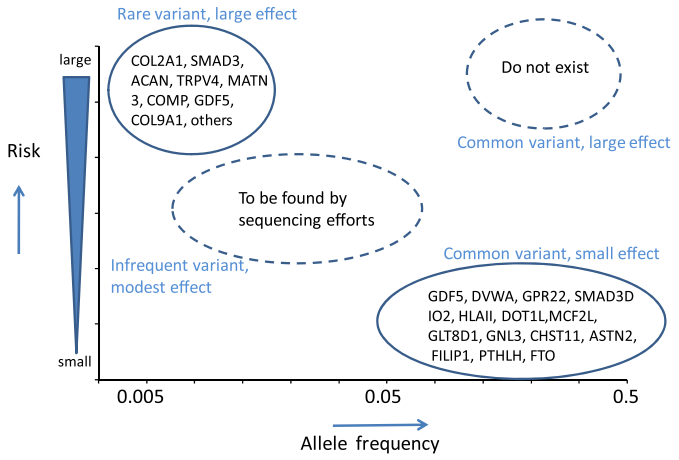


**Fig. 1.** Genetic architecture of susceptibility to OA. Genetic loci that are known to be associated to risk for OA comprise rare variants of large effect[3,4,10,11,40–42] and common variants of small effect[6,7,19,24,29,43–46]. Common variants of large effect have not been identified and probably do not exist. It is likely that uncommon variants of moderately large effect exist and contribute to risk for OA. These will most likely be found using next generation technology in families and population data.

often coincide with other major skeletal manifestations such as disproportionate short stature and skeletal malformations. These genetic skeletal disorders are caused by disturbances in the complex process of skeletal development at many stages and can vary a lot in their appearance and severity (see Ref. [1] for overview). These disorders are caused by mutations in genes encoding for proteins involved in various functions, i.e., extracellular matrix proteins and growth factors involved in cartilage/bone development, but also other less understood pathways such as post-translational processing and transport[2]. These "experiments of nature" can give us important insight into disease processes. Over the last year, several interesting new genes were identified to be the cause of early onset OA. First, mutations in *SMAD3* (*SMAD family member 3*) were identified as the cause of aneurysm-OA syndrome (AOS)[3]. This study showed the existence of coupling between aneurysms and OA via the TGF (Transforming Growth Factor) beta signaling pathway. A second gene identified to be involved in OA through family research is the *TRPV4* (*transient receptor potential cation channel, subfamily V, member 4*) gene[4]. Mutations that reduce channel activity were shown to cause inherited osteoarthropathy, which indicates a role for TRPV4 activity in articular cartilage homeostasis. Another interesting family was published in the journal *Nature Genetics* last year, where deletion of *MIR17HG*, encoding the miR-17∼92 polycistronic miRNA cluster on chromosome 13, resulted in a clear skeletal phenotype (microcephaly, short stature and digital abnormalities)[5]. These findings identified a regulatory function for miR-17∼92 in growth and skeletal development and represent the first example of a miRNA gene responsible for a syndromic developmental defect in humans. Interestingly, some of the genes found for the monogenic syndromic form of OA, have also been implicated to play a role in normal, late onset OA (also shown in Fig. 1). For example the *GDF5 (growth differentiation factor 5)* and *SMAD3* genes, for which it was shown that common polymorphisms in these genes are associated with knee and hip OA[6,7].

### *Late onset OA*

There have been several attempts to find OA genes *via* similar approaches in families where the common form of OA is segregating using the so-called linkage approach[8]. However, most linkage studies have been unsuccessful in identifying genes involved in OA, and so a switch to the genetic association study design has been made in OA research, similar to other common complex diseases. Until the technical possibility of examining the complete common genomic variation of the genome by GWAS technology became available, researchers had to choose their candidate genes based on prior knowledge. Since knowledge of the etiology of OA is poor and, in general, power of the single cohorts is low, these studies have yielded inconsistent results (extensively reviewed in Ref. [9]).

There is one well replicated risk gene which reaches the genome-wide significance level ($p < 5 \times 10^{-8}$) that originated from such a candidate gene study: the *GDF5* gene. Rare inherited skeletal dysplasias such as Chondrodysplasia grebe type and brachydactyly type C are caused by mutations in *GDF*[10,11]. Common functional variation in this gene, affecting transcription of this gene, was recently shown to be involved in late onset OA[12] in both Asians as well as Europeans[13,14].

## The GWAS era

GWAS have been successfully applied to the study of many complex diseases and in less than 5 years have identified more than 1,500 loci that predispose to diseases and quantitative traits (see www.genome.gov/gwastudies). Several GWAS studies for OA have been published over the last year[9]. It is evident that the number of published loci that were found robustly associated with OA up to the start of 2012 was low. The reason for this is – as we now know – most likely lack of power. Power in a genetic epidemiology study is dependent of sample size and phenotype heterogeneity.

### *Sample size*

The studies that had been published up to the start of 2012 all had low to moderate sample sizes (upto 2,300 cases and controls in the discovery phase). In other human diseases it is accepted that international collaboration is necessary to reach the required sample sizes. Examples of such large undertakings are the DIAGRAM consortium[15] (diabetes, >45,000 individuals), CARDIOGRAM/C4D consortium[16,17] (65,000 coronary artery disease cases and 130,000 controls) and the largest up to now: the GIANT-consortium which examines anthropometric traits in >250,000 individuals[18]. Within OA, there are several initiatives to collaborate in order to reach the sample sizes that are required to find new robust GWAS signals. The results of the largest collaborative effort up to now, the Arthritis Research UK (ARC)-sponsored effort arcOGEN in the UK collaborating with the European Union (EU)-sponsored TREAT-OA consortium, were presented at the OARSI-meeting in 2011 and 2012 and are about to be published at the time of writing[19]. With more than 7,000 OA cases in the discovery, they identified nine new genetic loci to be involved in OA. The large arcOGEN study shows that bigger sample sizes do result in larger number of genetic loci identified, and therefore it is anticipated that even larger collaborative efforts within the OA genetics field will result in more genes and novel pathways involved in OA, and consequently also a better understanding of the disease mechanism. During the last 9 months, several cohort studies have acquired new GWAS-data, including the OA Initiative, the Johnston County study, the Study of Osteoporotic Fractures and MrOs studies. These deeply phenotyped cohorts will certainly help genetics of OA forward.

### *Phenotypes in OA*

Figure 2 shows the number of identified genetic susceptibility loci (based on assessing only the common variants through GWAS)

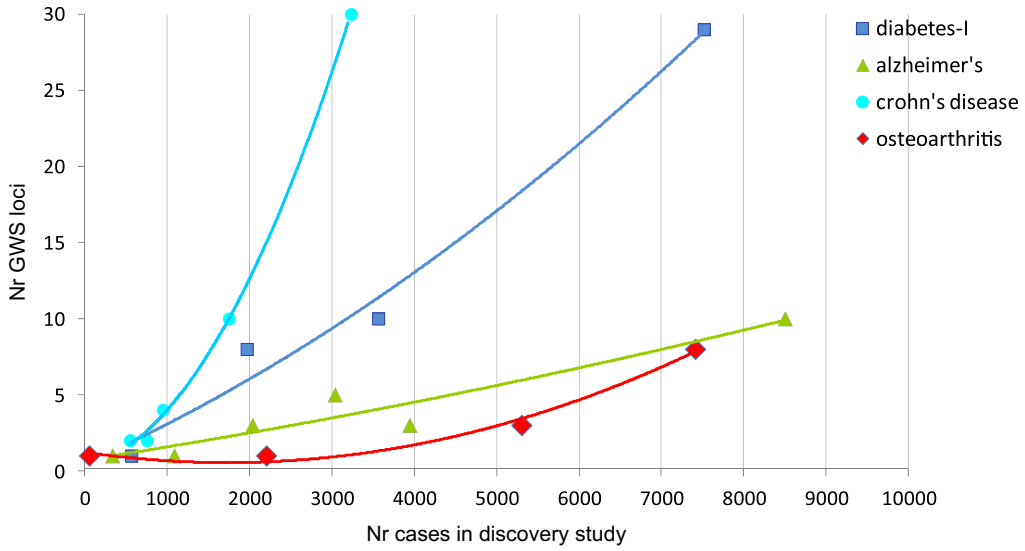


**Fig. 2.** Power of genetic studies depends on sample size and phenotype. The number of GWAS "hits" (SNPs with $p < 5 \times 10^{-8}$) as a function of sample size of the discovery cohort for several major complex diseases including OA.

as a function of the sample sizes of the original discovery GWAS for several dichotomous disease outcomes, including OA. It shows that there are large differences in number of identified loci between the different phenotypes, with Crohns' disease being a relatively easy trait where dozens of loci were found with relatively low sample size, while in Alzheimers' disease and OA only a modest number of loci were identified despite modestly large sample size. These differences in "genetic accessibility" could reflect the heterogeneity of the phenotype under study, which is an important determinant of power to identify robust association[20]. In contrast to many disorders that can be considered as the extreme of a normal distribution of a physiological parameter (such as osteoporosis is for bone mineral density (BMD)), OA is different, since it is not the extreme of a distribution of cartilage degradation. The diagnosis of OA is based on a combination of parameters including both clinical features (pain and stiffness) as well as a structural damage composite score (the most widely used is the Kellgren & Lawrence score, K&L score), which includes formation aspects of bone (osteophyte formation) and assessment of cartilage degradation. In general, it is suspected that OA is an extremely heterogeneous phenotype with differently defined disease status as an extra source of phenotype heterogeneity[21].

One might consider that for genetic studies, it can be worthwhile to examine so-called endophenotypes. Endophenotypes are measurable intermediate phenotypes that are generally closer to the action of the gene product than the disease status, and thus exhibit higher genetic signal-to-noise ratios[22]. Endophenotypes often provide much greater power to localize and identify disease-related QTLs than does disease status alone[23]. Useful endophenotypes in OA might be cartilage characteristics or joint shape among others. An example of this approach is a GWAS on cartilage thickness which was published by our group very recently, identifying the DOT1-like, histone H3 methyltransferase (*DOT1L*) gene to be involved in cartilage thickness, as measured by joint space width on a radiograph[24].

On the other hand, the major clinical outcome, pain, can also be considered as a disease phenotype on its own, with joint destruction as one of the risk factors for getting chronic joint pain (see also Fig. 3). Over the last year several papers were published that suggested involvement of several genes involved in OA-related pain. One study identified *via* a genome-wide linkage analysis in mice, the *P2X7* receptor gene to be involved in nerve-injury-induced pain behavior[25]. Subsequently, in the same study, a missense variant in the *P2RX7* was found to be associated with post-mastectomy pain

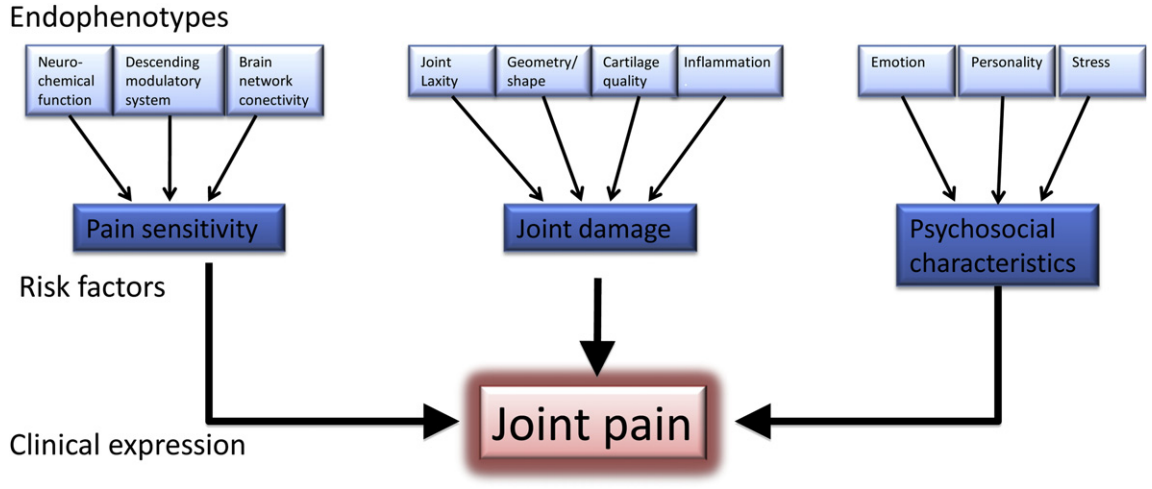


**Fig. 3.** Relationship among major and minor phenotypes and sub-phenotypes to be distinguished in OA research.

as well as OA-related pain. Another study showed a genetic association between variants in the gene that encodes PACE4, the proprotein convertase subtilisin/kexin type 6 (*PCSK6*) gene and OA-related pain[26]. Although both reported genetic associations need further replication, these studies indicate that pain can, and maybe should, be studied as a separate trait.

## Biological mechanism behind genetic associations

The exact pathophysiological mechanisms are largely unknown for the majority of the risk alleles of the polymorphisms identified *via* GWAS screening in OA. One of the problems is the identification of the culprit gene, among the several genes that lie within a GWAS hit region. Other problems can be the identification of the causative single nucleotide polymorphism (SNP) (among the many linked SNPs in such an area), and subsequently the unraveling of what the risk allele of the SNP actually does. For few genetic loci the functional risk allele has been elucidated. For example, the *GDF5* SNP was shown to exert its effect *via* regulation of transcription of the gene in which the risk allele led to lower expression levels[27]. Also SNPs near the *DIO2 (deiodinase, iodothyronine, type II)* gene were shown to affect transcription[28]. Both DIO2 and GDF5 are thought to play a role in skeletal development. In addition, the recently identified *DOT1L* gene was suggested to play a role in chondrogenic bone development *via* regulation of Wnt-signaling[24]. In contrast, it has been difficult to identify the underlying biological mechanism of one of the most robust genetic signals in OA: the Chr7q22 locus. This locus harbors six genes with no obvious candidate gene to pursue. Although there is some indication that *GPR22 (G-protein coupled receptor)* might be the causal gene, evidence is not conclusive[29], and also other genes are now being implicated[30,31]. Research is ongoing to elucidate the exact mechanism. For the nine newly identified loci from the arcoGEN study, some of them are intuitively easy to interpret, such as the *CHST11 (carbohydrate (chondroitin 4) sulfotransferase 11)* (involved in formation of chondroitin and dermatan sulfate, important components of cartilage proteoglycans) and *FTO (fat mass and obesity associated)* (increasing risk for knee OA by influencing BMI), but others need much more work.

## Next steps in OA genomic research

Genomic research in OA has progressed significantly over the last years. However, the resulting novel findings have not been accompanied by clinically applicable tools for risk prediction, diagnosis or therapeutic interventions. Novel methods and approaches are emerging that follow the recent genetic findings for OA. There are a few key areas that will be critical for successful translation of the genomic findings to the clinic. First, Next Generation Sequencing (NGS) will identify rare and novel variants associated with OA or its endophenotypes, beyond the ones that were captured by current GWAS techniques. Since GWAS assesses only 0.1% of the nucleotides in the human genome much progress can be expected from application of NGS (that assesses virtually all nucleotides) although the approach has yet to mature in terms of study designs and interpretation of data. Second, technological advances will also put forward the fields of transcriptomics and epigenetics as possible tools to improve understanding of the pathophysiology of OA and clinical translation.

### *NGS*

Sequencing technology has now advanced to a stage where it is possible to generate a complete catalog of all variants present within a given DNA sequence rather than having to rely on markers and patterns of linkage disequilibrium. NGS platforms have markedly decreased the cost of DNA sequencing when compared with Sanger sequencing. Hybridization approaches in NGS have enabled selection of the portion of the genome that is protein coding (roughly 1% of the genome), the so-called “exome”. Sequencing just the exome (rather than the entire genome) is more cost-effective and targets the part of the genome that is most likely to directly affect the translated proteins. Hence biological interpretation is relatively straightforward.

Over the past year, whole-exome sequencing has been successfully utilized to identify new genes for several Mendelian diseases that were not yet elucidated with the classical linkage approach[32]. Most studies have sequenced the exomes of one or a few individuals affected with the disorder. Variants seen in these individuals are typically compared with those from reference individuals (unaffected individuals who are family members or unrelated). Variants that are shared by affected individuals and are not present in the unaffected population are considered causal candidates. Important to note is that over 3,000 novel mutations are identified in each individual (irrespective of case/control status), and so filtering out the mutations that are not causative *via* family-information is essential in this approach. It is anticipated that this strategy will also identify new mutations in unresolved Mendelian syndromic forms of early onset OA.

What about exome or whole-genome sequencing to identify rare variants that confer a large effect on common, complex form of OA? The optimal study design for complex traits will depend on the frequency of the genetic variant(s) that are the source of the association signal. However, without the extra information of a familial design, sequencing in the open population will meet its statistical boundaries. Statistical analysis needs to account for type I error in the huge datasets that are produced, since preliminary analysis of the first population-based NGS exome-datasets show over 300,000 new low-frequent mutations in 500 individuals. Consequently, there is also a need to develop methods that can distinguish phenotype-relevant variants from commonly shared alleles. Therefore, NGS has up to now mainly been applied to so-called extreme phenotypes, representing the extremes of a phenotype distribution in the hope to increase the likelihood of finding rare variants associated with the disease phenotype. For OA, such an extreme phenotype could be comparing young persons with OA at multiple sites against old persons without OA. There are also efforts ongoing to completely sequence large cohorts that have OA data. This will undoubtedly result in novel findings in the OA field, such as an enormous amount of new DNA-variants, including SNPs and structural variants, small and large insertions/deletions. Yet, replication strategies and the required power are now being debated given that the rare variants themselves might not so much be the “replicable unit” rather than the affected gene. Coupling all these genetic variants to function (through transcriptomics and epigenetics) is one of the major tasks of genomic research over the next decade.

### *Epigenetics*

Epigenetics investigates the mechanisms of the control of gene expression that is apart from genetic variants. The most investigated epigenetic modification is DNA methylation but also histone modifications and miRNAs are epigenetic phenomena that are under investigation. Epigenetics might be crucial for understanding the molecular basis of complex diseases such as OA[33]. Although epigenetic modifications can be genetically determined, for example by genetic variants in CpG sites and hence can result in differential methylation, they can also be modified by environmental influences, such as diet[34]. DNA methylation can be

studied using a candidate gene approach. This has recently been applied successfully for *GDF5*, where it was demonstrated that the genetic effect of the rs143383 SNP on GDF5 expression is modulated epigenetically by DNA methylation[35]. In addition, similar approaches were presented for *DIO2* at the OARSI, also indicating that this gene is under control of both genetic as well as epigenetic differences[28,36].

DNA methylation can now be studied on a genome-wide scale by methylation arrays for which the latest arrays contain 450,000 CpG sites. The first studies are appearing now that identify differential methylation signatures for disease states. A first pilot study examining 27,000 CpG sites in OA cartilage specimens was presented at the OARSI-meeting by Fernandez-Tajes *et al.* They showed suggestive evidence for different subtypes of OA based on clustering of the methylation profiles[37].

### *Transcriptomics*

Similar to epigenetic marks, gene expression is determined by both genetic and environmental influences. The transcriptome is highly dynamic and therefore gene-expression analysis at a single time point reflects a snapshot of the cellular/tissue condition, but does not necessarily reflect biological variation over time. In contrast to the very stable DNA molecules, RNA is instable and therefore sample preparation is more difficult and crucial.

RNA-expression analysis can be used to examine whether the identified genetic association is explained by regulating RNA-levels of a nearby gene by simultaneously determining RNA-expression levels and genetic variation (see also Fig. 4). This has been done on a genome-wide scale, to identify so-called expression QTL (eQTL) analysis for peripheral blood and other tissues, such as liver and brain. In bone and joint tissues such a genome-wide eQTL analysis has not been established yet. However, for some recently identified OA genes, such as *GDF5, DIO2*, and dihydrouridine synthase 4-like (*DUS4L*), HMG-box transcription factor 1 (*HBP1*), component of oligomeric golgi complex 5 (*COG5*) (the last three genes are part of the Chr7 cluster) allele specific expression in joint tissues showed that indeed these genes were differentially expressed between the different genotype groups of the genetic variants[27,28,30].

Several studies have analyzed (independent of genotype) genome-wide gene-expression microarray data in human end-stage OA cartilage collected at the time of joint replacement. These studies generally lack large numbers and independent replication of the findings and so replication of such findings is also highly recommended before interpretation and validity can be properly assessed. Expression signatures can also be used as biomarkers for disease. This is actively done in the clinical practice of oncology, but not for OA up to now. Of course, the target tissues of the joint cannot be sampled for biomarker use in early disease, and therefore use of transcriptomics as biomarkers should be limited to easily obtainable tissues, such as blood. An interesting first study of Attur *et al.* showed an association between circulating Interleukine 1 (IL-1)beta expression levels and progression of knee OA[38], which supports the idea of low grade systemic inflammation being present in OA. This finding, of course needs further large scale replication and validation, before it can considered for implementation in the clinical practice.

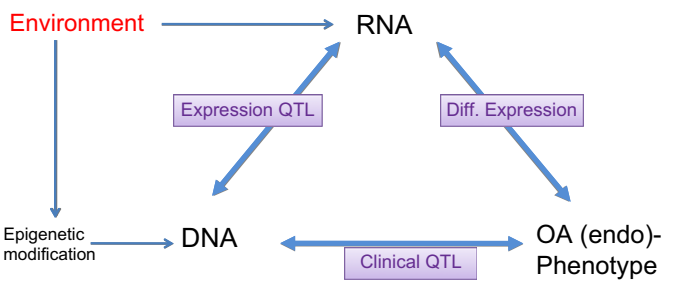


**Fig. 4.** Ongoing efforts in genetics and genomics to disentangle the functional genome and elucidate the molecular mechanisms underlying OA.

New technological advances are also entering the transcriptomics research field. With the advent of NGS technologies, RNA sequencing (RNA-seq) by NGS has emerged as a powerful tool for transcriptome analysis.

It is likely that the microarray-based gene-expression profiling technology will be replaced by RNA-seq based expression profiling, because RNA-seq will make it possible to explore much more variation of the transcriptome compared to array-technology. Besides the much wider dynamic range of signal, RNA-seq has the ability to comprehensively detect novel transcripts and mRNA variants resulting from alternative promoter usages, splicing, polyadenylation and sequence variation; and lowered background.

At the OARSI-meeting in Barcelona, a first study using this RNA-seq method was presented which highlighted some new genes upregulated in OA cartilage compared to controls[39]. It is clear that these new approaches render an enormous amount of data, and new methods to analyze these rich data are needed and being developed.

## Summary

Technological advances in analyzing the major biomolecules DNA and RNA, together with better statistical modeling and improved computational capacity and software have transformed the genomics field rapidly over the last years. Genetic research in particular has received an enormous boost when GWAS was introduced as a technique to identify genes in complex traits. For OA, GWAS is now becoming more successful in identifying new OA genes. More power in genetic analysis, reached *via* a combination of increasing sample size and refinement of the OA-phenotypes will continue this success. The next technological step is coming with NGS, making it possible to fully sequence many genomes at the basepair level. Developing in parallel are a number of new approaches at different levels: epigenomics, transcriptomics, and maybe also proteomics and metabolomics (although not reviewed here). These new genome-wide approaches can offer us new insights in genome function. Studying the DNA sequence with GWAS and NGS at the population level will result in a nearly complete picture of genetic variation in the static DNA-backbone in relation to the many phenotypes and clinical states we can distinguish in OA. An important next step is to add dynamic data of epigenomics and transcriptomics. Integrating these areas is necessary to fully understand a complex disease, like OA. The next few years will yield more insights into the relationship between genotype and OA-phenotypes, and hopefully identify dominant gene networks and pathways in patient subsets of OA. This would then open up new avenues for targeted therapy.

## Author contributions

Joyce BJ van Meurs contributed to the conception and design, collection and assembly of data, analysis and interpretation of data, drafting/revisions of article, as well as final approval of the article.

Andre G Uitterlinden contributed to the conception and design, analysis and interpretation of data, drafting/revisions of article, as well as final approval of the article.

**Role of funding source**
None of the authors received any funding related to the writing of this manuscript.

**Conflicts of interest**
JvM and AU have no conflict of interest to disclose.

**Acknowledgments**

Our research is funded by the European Commission (HEALTH-F2-2008-201865, GEFOS; HEALTH-F2-2008-35627, TREAT-OA), Netherlands Organisation for Scientific Research (NWO) (VIDI grant 917103521), the Netherlands Consortium for Healthy Aging (NCHA), and the Netherlands Genomics Initiative (NGI). The funding bodies had no role in the writing or the content of this manuscript.

**References**

1. Warman ML, Cormier-Daire V, Hall C, Krakow D, Lacman R, LeMerrer M, *et al*. Nosology and classification of genetic skeletal disorders: 2010 revision. Am J Med Genet A 2011;155A: 943–68.
2. Cormier-Daire V. Spondylo-epi-metaphyseal dysplasia. Best Pract Res Clin Rheumatol 2008;22:33–44.
3. van de Laar IMBH, Oldenburg RA, Pals G, Roos-Hesselink JW, de Graaf BM, Verhagen JMA, *et al*. Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. Nat Genet 2011;43:121–6.
4. Lamande SR, Yuan Y, Gresshoff IL, Rowley L, Bellupccio D, Kaluarachchi K, *et al*. Mutations in TRPV4 cause an inherited arthropathy of hands and feet. Nat Genet 2011;43:1142–6.
5. de Pontual L, Yao E, Callier P, Faivre L, Drouin V, Cariou S, *et al*. Germline deletion of the miR-17 approximately 92 cluster causes skeletal and growth defects in humans. Nat Genet 2011;43:1026–30.
6. Valdes AM, Spector TD, Tamm A, Kisand K, Doherty SA, Dennison EM, *et al*. Genetic variation in the SMAD3 gene is associated with hip and knee osteoarthritis. Arthritis Rheum 2010;62:2347–52.
7. Evangelou E, Chapman K, Meulenbelt I, Karassa FB, Loughlin J, Carr A, *et al*. Large-scale analysis of association between GDF5 and FRZB variants and osteoarthritis of the hip, knee, and hand. Arthritis Rheum 2009;60:1710–21.
8. Lee YH, Rho YH, Choi SJ, Ji JD, Song GG. Osteoarthritis susceptibility loci defined by genome scan meta-analysis. Rheumatol Int 2006;26:959–63.
9. Valdes AM, Spector TD. Genetic epidemiology of hip and knee osteoarthritis. Nat Rev Rheumatol Jan 2011;7(1):23–32.
10. Thomas JT, Kilpatrick MW, Lin K, Erlacher L, Lembessis P, Costa T, *et al*. Disruption of human limb morphogenesis by a dominant negative mutation in CDMP1. Nat Genet 1997;17: 58–64.
11. Polinkovsky A, Robin NH, Thomas JT, Irons M, Lynn A, Goodman FR, *et al*. Mutations in CDMP1 cause autosomal dominant brachydactyly type C. Nat Genet 1997;17:18–9.
12. Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, Saito S, *et al*. A functional polymorphism in the 5′ UTR of GDF5 is associated with susceptibility to osteoarthritis. Nat Genet 2007;39: 529–33.
13. Vaes RB, Rivadeneira F, Kerkhof JM, Hofman A, Pols HA, Uitterlinden AG, *et al*. Genetic variation in the GDF5 region is associated with osteoarthritis, height, hip axis length and fracture risk: the Rotterdam study. Ann Rheum Dis 2009;68: 1754–60.
14. Valdes AM, Evangelou E, Kerkhof HJ, Tamm A, Doherty SA, Kisand K, *et al*. The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. Ann Rheum Dis 2011;70:873–5.
15. Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, *et al*. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. Nat Genet 2012;44:659–69.
16. Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat Genet 2011;43:339–44.
17. Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, *et al*. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43:333–8.
18. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, *et al*. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 2010;467:832–8.
19. arcOGEN Consortium and arcOGEN Collaborators. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. Lancet 2012 July 2 [epub ahead of print].
20. Ioannidis JP, Thomas G, Daly MJ. Validating, augmenting and refining genome-wide association signals. Nat Rev Genet 2009;10:318–29.
21. Kerkhof HJ, Meulenbelt I, Akune T, Arden NK, Aromaa A, Bierma-Zeinstra SM, *et al*. Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. Osteoarthritis Cartilage 2011;19:254–64.
22. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 2003;160:636–45.
23. Blangero J, Williams JT, Almasy L. Novel family-based approaches to genetic risk in thrombosis. J Thromb Haemost 2003;1:1391–7.
24. Castano Betancourt MC, Cailotto F, Kerkhof HJ, Cornelis FM, Doherty SA, Hart DJ, *et al*. Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. Proc Natl Acad Sci USA 2012;109:8218–23.
25. Sorge RE, Trang T, Dorfman R, Smith SB, Beggs S, Ritchie J, *et al*. Genetically determined P2X7 receptor pore formation regulates variability in chronic pain sensitivity. Nat Med 2012;18: 595–9.
26. Malfait AM, Seymour AB, Gao F, Tortorella MD, Le Graverand-Gastineau MP, Wood LS, *et al*. A role for PACE4 in osteoarthritis pain: evidence from human genetic association and null mutant phenotype. Ann Rheum Dis 2012;71:1042–8.
27. Egli RJ, Southam L, Wilkins JM, Lorenzen I, Pombo-Suarez M, Gonzalez A, *et al*. Functional analysis of the osteoarthritis susceptibility-associated GDF5 regulatory polymorphism. Arthritis Rheum 2009;60:2055–64.
28. Bos SD, Bovee JV, Duijnisveld BJ, Raine EV, van Dalen WJ, Ramos YF, *et al*. Increased type II deiodinase protein in OA-affected cartilage and allelic imbalance of OA risk polymorphism rs225014 at DIO2 in human OA joint tissues. Ann Rheum Dis 2012;71:1254–8.
29. Kerkhof HJ, Lories RJ, Meulenbelt I, Jonsdottor I, Valdes AM, Arp P, *et al*. A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22. Arthritis Rheum 2010;62:499–510.
30. Raine EV, Wreglesworth N, Dodd AW, Reynard LN, Loughlin J. Gene expression analysis reveals HBP1 as a key target for the

osteoarthritis susceptibility locus that maps to chromosome 7q22. Ann Rheum Dis 2012 Jul 6 [Epub ahead of print].
31. Cailotto F, Luyten FP, Lories RJ. Cog5 inhibition induces glycosylation defects affecting chondrogenesis and interfering with Wnt, but not Bmp signaling. Osteoarthritis Cartilage Apr 2012;20:S142–3.
32. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, *et al*. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 2011;12:745–55.
33. Petronis A. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. Nature 2010;465:721–7.
34. Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and implications. Nat Rev Genet 2011;13:97–109.
35. Reynard LN, Bui C, Canty-Laird EG, Young DA, Loughlin J. Expression of the osteoarthritis-associated gene GDF5 is modulated epigenetically by DNA methylation. Hum Mol Genet 2011;20:3450–60.
36. den Hollander W, Bos SD, Ramos YF, Lakenberg N, van der Breggen R, Duijnisveld BJ, *et al*. Cpg sites of osteoarthritis susceptibility gene Dio2 are differentially methylated in arthritic compared to preserved cartilage. Osteoarthritis Cartilage 2012;20:S196.
37. Fernandez-Tajes J, Fernandez-Moreno M, Tamayo-Novas M, Mosquera-Rey A, Oreira-Villar N, Fernandez-Moreno M, *et al*. The analysis of the genome-wide DNA methylation profile of human articular chondrocytes reveals different forms of OA. Osteoarthritis Cartilage 2012;20:S42.
38. Attur M, Belitskaya-Levy I, Oh C, Krasnokutsky S, Greenberg J, Samuels J, *et al*. Increased interleukin-1beta gene expression in peripheral blood leukocytes is associated with increased pain and predicts risk for progression of symptomatic knee osteoarthritis. Arthritis Rheum 2011;63:1908–17.
39. Xu Y, Barter MJ, Swan DC, Rankin KS, Rowan AD, Santibanez-Koref M, *et al*. Comparison of osteoarthritis and normal hip cartilage transcriptomes using Rna-Seq reveals new candidate gene targets and associated pathways. Osteoarthritis Cartilage 2012;20:S43.
40. Husar-Memmer E, Ekici A, Al Kaissi A, Sticht H, Mager B, Schett G, *et al*. Premature osteoarthritis as presenting sign of type II collagenopathy: a case report and literature review. Semin Arthritis Rheum June 2012;2012 [epub ahead of print].
41. Gleghorn L, Ramesar R, Beighton P, Wallis G. A mutation in the variable repeat region of the aggrecan gene (AGC1) causes a form of spondyloepiphyseal dysplasia associated with severe, premature osteoarthritis. Am J Hum Genet 2005;77:484–90.
42. Jackson GC, Mittaz-Crettol L, Taylor JA, Mortier GR, Spranger J, Zabel B, *et al*. Pseudoachondroplasia and multiple epiphyseal dysplasia: a 7-year comprehensive analysis of the known disease genes identify novel and recurrent mutations and provides an accurate assessment of their relative contribution. Hum Mutat 2012;33:144–57.
43. Meulenbelt I, Min JL, Bos S, Riyazi N, Houwing-Duistermaat JJ, van der Wijk HJ, *et al*. Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. Hum Mol Genet 2008;17:1867–75.
44. Day-Williams AG, Southam L, Panoutsopoulou K, Rayner NW, Esko T, Estrada K, *et al*. A variant in MCF2L is associated with osteoarthritis. Am J Hum Genet 2011;89:446–50.
45. Miyamoto Y, Shi D, Nakajima M, Ozaki K, Sudo A, Kotani A, *et al*. Common variants in DVWA on chromosome 3p24.3 are associated with susceptibility to knee osteoarthritis. Nat Genet 2008;40:994–8.
46. Nakajima M, Takahashi A, Kou I, Rodriguez-Fontenla C, Gomez-Reino JJ, Furuichi T, *et al*. New sequence variants in HLA class II/III region associated with susceptibility to knee osteoarthritis identified by genome-wide association study. PLoS One 2010;5:e9723.